A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex

Saskia E. J. de Vries1,5*, Jerome A. Lecoq1,5*, Michael A. Buice1,5*, Peter A. Groblewski1, Gabriel K. Ocker1, Michael Oliver1, David Feng1, Nicholas Cain1, Peter Ledochowitsch1, Daniel Millman1, Kate Roll1, Marina Garrett1, Tom Keenan1, Leonard Kuan1, Stefan Mihalas1, Shawn Olsen1, Carol Thompson1, Wayne Wakeman1, Jack Waters1, Derric Williams1, Chris Barber1, Nathan Berbesque1, Brandon Blanchard1, Nicholas Bowles1, Shiella D. Caldejon1, Linzy Casal1, Andrew Cho1, Sissy Cross1, Chinh Dang1, Tim Dolbeare1, Melissa Edwards1, John Galbraith1, Nathalie Gaudreauult1, Terri L. Gilbert1, Fiona Griffin1, Perry Hargrave1, Robert Howard1, Lawrence Huang1, Sean Jewell2, Nika Keller1, Ulf Knoblich1, Josh D. Larkin1, Rachael Larsen1, Chris Lau1, Eric Lee2, Felix Lee1, Arielle Leon1, Lu Li1, Fuhui Long1, Jennifer Luviano1, Kyla Mace1, Thuyanh Nguyen1, Jed Perkins1, Miranda Robertson1, Sam Seid1, Eric Shea-Brown1,3, Jianghong Shi3, Nathan Sjoquist1, Cliff Slaughterbeck1, David Sullivan1, Ryan Valenza1, Casey White1, Ali Williford1, Daniela M. Witten1,2,4, Jun Zhuang1, Hongkui Zeng1, Colin Farrell1, Lydia Ng1, Amy Bernard1, John W. Phillips1, R. Clay Reid1 and Christof Koch1

To understand how the brain processes sensory information to guide behavior, we must know how stimulus representations are transformed throughout the visual cortex. Here we report an open, large-scale physiological survey of activity in the awake mouse visual cortex: the Allen Brain Observatory Visual Coding dataset. This publicly available dataset includes the cortical activity of nearly 60,000 neurons from six visual areas, four layers, and 12 transgenic mouse lines in a total of 243 adult mice, in response to a systematic set of visual stimuli. We classify neurons on the basis of joint reliabilities to multiple stimuli and validate this functional classification with models of visual responses. While most classes are characterized by responses to specific subsets of the stimuli, the largest class is not reliably responsive to any of the stimuli and becomes progressively larger in higher visual areas. These classes reveal a functional organization wherein putative dorsal areas show specialization for visual motion signals.

Traditional understanding, based on decades of research, is that visual cortical activity can be largely characterized by responses to a specific set of local features (modeled with linear filters followed by nonlinearities) and that these features become more selective and specialized in higher cortical areas1–4. However, it remains unclear to what extent this understanding can account for the whole of V1 (refs. 5–7), let alone the rest of the visual cortex. A key challenge results from the fact that this understanding is based on many small studies, recording responses from different stages in the circuit and using different stimuli and analyses5. The inherent experimental selection biases and lack of standardization of this approach introduce additional obstacles to creating a cohesive understanding of cortical function. On the basis of these issues, influential reviews have questioned the validity of this standard model5–7 and have argued that “what would be most helpful is to accumulate a database of single-unit or multi-unit data (stimuli and neural responses) that would allow modelers to test their best theory under ecological conditions” (ref. 5). To address these issues, we conducted a survey of visual responses across multiple layers and areas in the awake mouse visual cortex, while using a diverse set of visual stimuli. This survey was executed in pipeline fashion, with standardized equipment and protocols and with strict quality-control measures not dependent upon stimulus-driven activity (Methods).

Previous work in mouse has revealed functional differences among cortical areas in layer 2/3 in terms of the spatial and temporal frequency tuning of neurons in each area6,9. However, it is not clear how these differences extend across layers and across diverse neuron populations. Here we expand such functional studies to include 12 transgenically defined neuron populations, including Cre driver

1Allen Institute for Brain Science, Seattle, WA, USA. 2Department of Statistics, University of Washington, Seattle, WA, USA. 3Department of Applied Mathematics, University of Washington, Seattle, WA, USA. 4Department of Biostatistics, University of Washington, Seattle, WA, USA. 5These authors contributed equally: Saskia E. J. de Vries, Jerome A. Lecoq, Michael A. Buice. *e-mail: saskiad@alleninstitute.org; jeromel@alleninstitute.org; michaelbu@alleninstitute.org
lines for excitatory populations across four cortical layers (from layer 2/3 to layer 6) and for two inhibitory populations (defined by Vip and Sst expression). Further, it is known that stimulus statistics affect visual responses, such that responses to natural scenes cannot be well predicted by responses to noise or grating stimuli. To examine the extent of this discrepancy and its variation across areas and layers, we designed a stimulus set that included both artificial (gratings and noise) and natural (scenes and movies) stimuli. While artificial stimuli can be easily parameterized and interpreted, natural stimuli are closer to what is ethologically relevant. Finally, as recording modalities have enabled recordings from larger populations of neurons, it has become clear that populations might code visual and behavioral activity in a way that is not apparent by considering single neurons alone. Here we imaged populations of neurons (173±115 neurons for excitatory populations and 19±11 neurons for inhibitory populations, mean±s.d.) to explore both single-neuron and population coding properties.

We find that 77% of neurons in the mouse visual cortex respond to at least one of these visual stimuli, with many showing classical tuning properties, such as orientation- and direction-selective responses to gratings. These tuning properties exhibit differences across cortical areas and Cre lines. While subtle differences do exist between the excitatory Cre lines, these populations of excitatory neurons are largely similar; the more marked differences are among the inhibitory interneurons. The responses to all stimuli are highly sparse and variable. We find that the variability of responses is not strongly correlated across stimuli, in general, but this variability provides evidence for functional response classes. We validate these functional response classes with a model of neural activity that contains most of the basic features found in visual neurophysiological modeling (for example, ‘simple’ and ‘complex’ components) as well as the running speed of the mouse. For one class of neurons, these models perform quite well, predicting responses to both artificial and natural stimuli equally well. However, for many neurons, the models provide a poor description, particularly for those in our largest single class of neurons, those that respond reliably to none of our visual stimuli. The representation of these response classes across areas reveals a separation of motion processing from spatial selectivity across areas.

**Results**

By using adult C57BL/6 mice (age 108±16 d, mean±s.d.) that expressed a genetically encoded calcium sensor (GCaMP6f) under the control of specific Cre driver lines (ten excitatory lines and two inhibitory lines), we imaged the activity of neurons in response to a battery of diverse visual stimuli. Data were collected from six different cortical visual areas (V1, LM, AL, PM, AM and RL) and four different cortical layers. Visual responses of neurons at the retinotopic center of gaze were recorded in response to drifting gratings (DG), flashed static gratings (SG), locally sparse noise, natural scenes (NS), and natural movies (NM) (Fig. 1f), while the mouse was awake and free to run on a rotating disc. In total, 59,610 neurons were imaged from 432 experiments (Table 1). Each experiment consisted of three 1-h imaging sessions, with 33.6% of neurons matched across all three sessions; the rest were present in either one or two sessions (Methods).

To systematically collect physiological data on this scale, we built data collection and processing pipelines (Fig. 1). The data collection workflow progressed from surgical headpost implantation and craniotomy to retinotopic mapping of cortical areas using intrinsic signal imaging, in vivo two-photon calcium imaging of neuronal activity, brain fixation and histology using serial two-photon tomography (Fig. 1a–c). To maximize data standardization across experiments, we developed multiple hardware and software tools (Fig. 1d). One of the key components was the development of a registered coordinate system that allowed an animal to move from one data collection step to the next, on different experimental platforms, and maintain the same experimental and brain coordinate geometry (Methods). In addition to such hardware instrumentation, formalized standard operating procedures and quality-control metrics were crucial for the collection of these data over several years (Fig. 1e).

Following data collection, fluorescence movies were processed using automated algorithms to identify somatic regions of interest (ROIs) (Methods). Segmented ROIs were matched across imaging sessions. For each ROI, events were detected from $\Delta F/F$ by using an L0-regularized algorithm (Methods). The median average event magnitude during spontaneous activity was 0.0004 (arbitrary units; AU; event magnitude has the same units as $\Delta F/F$) and showed some dependence on depth and on transgenic Cre line (Extended Data Fig. 1).

For each neuron, we computed the mean response to each stimulus condition using the detected events, and we parameterized the neuron’s tuning properties. Many neurons showed robust responses, exhibiting orientation-selective responses to gratings, localized spatial receptive fields, and reliable responses to natural scenes and movies (Fig. 2a–f and Extended Data Fig. 2). For each neuron and each categorical stimulus (that is, drifting gratings, static gratings, and natural scenes), the preferred stimulus condition was identified as the one which evoked the largest mean response for that stimulus (for example, the orientation and temporal frequency with the largest mean response for drifting gratings). For each trial, the activity of a neuron was compared to a distribution of activity for that neuron taken during the epoch of spontaneous activity and a P value was computed. If at least 25% of the trials of the neuron’s preferred condition had a significant difference from the distribution of spontaneous activities ($P<0.05$), the neuron was labeled responsive to that stimulus (the Methods describes the responsiveness criteria for locally sparse noise and natural movies). Neurons meeting this criterion showed a change in activity with some degree of reproducibility across trials. The maximum evoked responses were an order of magnitude larger than the spontaneous activity (Extended Data Fig. 1; median of 0.006 (AU) for neurons responsive to drifting gratings).

In total, 77% of neurons were responsive to at least one of the visual stimuli presented (Fig. 2g). The percentage of responsive neurons depended on area and stimulus, such that V1 and LM showed the highest number of visually responsive neurons. This proportion dropped in other higher visual areas and was lowest in RL, where only 33% of neurons responded to any of the stimuli. Natural movies elicited responses from the most neurons, while static gratings elicited responses from the fewest (Fig. 2h). In addition to varying by area, the percentage of responsive neurons also varied by Cre line and layer, suggesting functional differences across these dimensions (Extended Data Figs. 3–7).

For responsive neurons, visual responses were parameterized by computing several metrics, including preferred spatial frequency, preferred temporal frequency, direction selectivity, receptive field size, and lifetime sparseness (Methods). We mapped these properties across cortical areas, layers, and Cre lines to examine the functional differences across these dimensions (Fig. 3 and Supplementary Figs. 1 and 2).

Comparisons across areas and layers revealed that direction selectivity was highest in layer 4 of V1 (Fig. 3b). While previous literature has found higher direction selectivity in layer 4 within V1 (ref. 18), we found here that this result was significant across all layer 4-specific Cre lines and extended to the higher visual areas as well. Comparisons across the higher visual areas revealed that, in superficial layers, the lateral higher visual areas (LM and AL) showed significantly higher direction selectivity than the medial ones (PM and AM), but this difference was not significant in the deeper layers.
This erosion of the differences between higher visual areas in deeper layers was found for all metrics reported here, with the population differences less pronounced and often not significant in layers 5 and 6 (Fig. 3c–e and Supplementary Fig. 2).

Across all areas, layers, and stimuli, visual responses in mouse cortex were highly sparse (Fig. 3f). When considering the responses to natural scenes, we found that most neurons responded to very few scenes (examples in Fig. 2d). The sparseness of individual neurons was measured by using lifetime sparseness, which captures the selectivity of a neuron's mean response to different stimulus conditions\(^{19,20}\) (Methods). A neuron that responds strongly to only a few scenes will have a lifetime sparseness close to 1 (Supplementary Fig. 3), whereas a neuron that responds broadly to many scenes will have a lower lifetime sparseness. Excitatory neurons had a median lifetime sparseness of 0.77 in response to natural scenes. While Sst neurons were comparable to excitatory neurons (median 0.77), Vip neurons exhibited low lifetime sparseness (median 0.36). Outside of layer 2/3, there was lower lifetime sparseness in areas RL, AM,
and PM than in areas V1, LM, and AL. Lifetime sparseness did not increase outside of V1; responses did not become more selective in the higher visual areas (Fig. 3I and Supplementary Fig. 3).

The pattern in single-neuron direction selectivity was reflected in our ability to decode the visual stimulus from single-trial population vector responses, by using all neurons, responsive and unresponsive (Fig. 4a and Supplementary Fig. 4). We used a k-nearest-neighbors classifier to predict the grating direction. When matching the tuning properties, areas V1, AL, and LM showed higher decoding performance than areas AM, PM, and RL, and these differences were more pronounced in superficial layers than in deeper layers. Similarly, the population sparseness (Supplementary Fig. 3), a measure of the selectivity of each scene (that is, how many neurons respond on a given trial), largely mirrored the high average lifetime sparseness of the underlying populations (Fig. 4b). Such high sparseness suggests that neurons are active at different times and thus their activities are weakly correlated. The noise correlations of the populations reflected the results on population sparsity, where excitatory populations showed weak correlations (median 0.02) while inhibitory populations showed somewhat higher correlations (Sst neurons, median 0.06; Vip neurons, median 0.15) (Fig. 4c). The structure of the correlations in each population may serve to either help or hinder information processing16,21. To test this, we measured the decoding performance when stimulus trials were shuffled to break trial-wise correlations. This had variable effects on decoding the decoding performance when stimulus trials were shuffled to cortex showed large trial-to-trial variability. Even when removing grating stimulus.

Correlations among Sst neurons were informative about the drifting properties, areas V1, AL, and LM showed higher decoding performance for Sst populations, on the other hand, was more consistently hurt by removing correlations, suggesting that the high correlations among Sst neurons were informative about the drifting grating stimulus.

For all stimuli, the visually evoked responses throughout the cortex showed large trial-to-trial variability. Even when removing the neurons deemed unresponsive, the percentage of responsive trials for most responsive neurons at their preferred conditions was low, with the median less than 50% (Fig. 5a and Supplementary Fig. 5). This means that the majority of neurons in the mouse visual cortex do not usually respond to individual trials, even when presented with the stimulus condition that elicits their largest average response. This was true throughout the visual cortex, although V1 showed slightly more reliable responses than higher visual areas and Sst interneurons, in particular, showed very reliable responses. The variability of responses was reflected in the high coefficient of variation, with median values for excitatory neurons above 2, indicating that these neurons are super-Poisson (Fig. 5b). We sought to capture this variability with a simple categorical model for drifting grating responses that attempts to predict the trial response (the integrated event magnitude during each trial) from the stimulus condition (the direction and temporal frequency of the grating or whether the trial was a blank sweep). This regression quantifies how well the average tuning curve predicts the response for each trial. Comparing the trial responses to the mean tuning curve showed a degree of variability even when the model was fairly successful (Fig. 5c). In line with this variability in visual responses, this model did a poor job of predicting responses to drifting gratings for most neurons (Fig. 5d). Few neurons were well predicted by their average tuning curve alone (21% of responsive neurons had \( r > 0.5 \), which became 11% when considering all neurons, where \( r \) is the cross-validated correlation between model prediction and actual response). As expected, the ability to predict responses was correlated with the measured variability (\( r = 0.8 \), Pearson correlation).

One possible source of trial-to-trial variability is the locomotor activity of the mouse. Previous studies have shown that, not only is some neural activity in the mouse visual cortex associated with running, but visual responses are also modulated by running22–26. The mice in our experiments were free to run on a disc, and animals showed a range of running behaviors (Supplementary Fig. 6). When ignoring the stimulus, we found that the activities of some neurons

| Table 1 | Visual coding dataset |
|---------|-------------------------|
| Cre line          | Layers | E/I | n (M/F) | Age range (d) | V1 | LM | AL | PM | AM | RL |
| Emx1-IRES-Cre;Camk2a-tTA;Ai93 | 2/3, 4, 5 | E  | 18 (13/5) | 73-156 | 3,073 (10) | 2,098 (8) | 1,787 (7) | 835 (4) | 457 (3) | 2,152 (9) |
| Slc17a7-IRES2-Cre;Camk2a-tTA;Ai93 | 2/3, 4, 5 | E  | 31 (20/11) | 80-149 | 4,840 (17) | 3,230 (16) | 374 (2) | 1,970 (15) | 235 (2) | 137 (2) |
| Cux2-CreERT2;Camk2-tTA;Ai93 | 2/3, 4 | E  | 38 (26/12) | 79-155 | 5,081 (16) | 2,792 (11) | 3,103 (13) | 2,361 (13) | 1,616 (11) | 1,578 (12) |
| Rorb-IRES2-Cre;Camk2a-tTA;Ai93 | 4 | E  | 24 (14/10) | 77-141 | 2,218 (8) | 1,191 (6) | 1,242 (6) | 764 (7) | 735 (8) | 1,126 (5) |
| Scnn1a-Tg3-Cre;Camk2a-tTA;Ai93 | 4 | E  | 7 (3/4) | 75-133 | 1,873 (9) |
| Nr5a1-Cre;Camk2a-tTA;Ai93 | 4 | E  | 23 (15/8) | 78-168 | 578 (8) | 421 (6) | 220 (6) | 331 (7) | 171 (6) | 1,354 (6) |
| Rbp4-Cre, KL100;Camk2a-tTA;Ai93 | 5 | E  | 23 (11/12) | 68-144 | 458 (7) | 485 (7) | 441 (6) | 509 (6) | 355 (8) | 93 (4) |
| Fezf2-CreER;Ai148 (corticofugal) | 5 | E  | 8 (4/4) | 88-134 | 407 (4) | 981 (5) |
| Tlx3-Cre, PL56;Ai148 (cortico-cortical) | 5 | E  | 7 (5/2) | 74-136 | 1181 (6) | 946 (3) |
| Ntsr1-Cre, GN220, Ai148 | 6 | E  | 10 (5/5) | 79-134 | 573 (6) | 719 (7) | 581 (5) |
| Sst-IRES-Cre, Ai148 | 4, 5 | I  | 30 (20/10) | 67-154 | 266 (17) | 301 (15) | 24 (1) | 247 (14) | 46 (2) |
| Vip-IRES-Cre, Ai148 | 2/3, 4 | I  | 24 (7/17) | 81-148 | 352 (17) | 315 (17) | 387 (16) |

The numbers of cells (and experiments) imaged for each Cre line in each cortical visual area are indicated. In total, 59,610 cells imaged in 432 experiments in 243 mice are included in this dataset. E, excitatory; I, inhibitory; F, female; M, male.
Neurons exhibit diverse responses to visual stimuli. Activity for four example neurons, including two excitatory neurons (Rorb, layer 4; Rbp4, layer 5) and two inhibitory neurons (Sst, layer 4; Vip, layer 2/3). ΔF/F (top, blue) and extracted events (bottom, black) are shown for each neuron. Star plots summarizing orientation and temporal frequency tuning for responses to the drifting grating stimulus. The arms of the star represent the different grating directions, and the rings represent the different temporal frequencies. At each intersection, the color of the circles corresponds to the strength of the response during a single trial of that direction and temporal frequency. (For details on response visualizations, see Extended Data Fig. 2.) Fan plots summarizing orientation and spatial frequency tuning for responses to static gratings. The arms of the fan represent the different orientations, and the arcs represent the spatial frequencies. For each condition, four phases of gratings were presented. cpd, cycles per degree. Corona plots summarizing responses to natural scenes. Each arm represents the response to an image, with individual trials represented by circles whose color corresponds to the strength of the response on that trial. Track plots summarizing responses to natural movies. The response is represented as a raster plot moving clockwise around the circle. Ten trials are represented in red, along with the mean peristimulus time histogram (PSTH) in the outer ring in blue. Receptive field subfields mapped using locally sparse noise. Percentage of neurons that responded to at least one stimulus across cortical areas. Colors correspond to the labels at the top of the figure. See Extended Data Figs. 3–7 for sample sizes.

were correlated with the running speed (Fig. 5c). While layer 5 showed strong correlations in all visual areas, in the other layers, V1 had stronger correlations than the higher visual areas, with some visual areas showing median negative correlations. Within V1, the inhibitory interneurons showed the strongest correlation with running, most notably Vip neurons in layer 2/3 (median 0.25), while the excitatory neurons showed weaker correlations (median 0.03).

For experiments with sufficient numbers of stimulus trials for a neuron’s preferred condition when the mouse was both stationary and running (>10% trials for each), we compared the responses in these two states. In line with other reports, many neurons showed modulated responses, but the effect was modest (Fig. 5f). The majority of neurons showed enhanced responses. When considering the entire population, there was a 1.9-fold increase in the median evoked response with running. The effect on individual neurons, however, was varied such that only 13% of neurons showed significant modulation in these conditions (P < 0.05, Kolmogorov–Smirnov test).

To test whether running accounted for the variability in trial-wise responses to visual stimuli, we included a binary running state as a condition-dependent gain into the categorical regression (that is, computing separate tuning curves for the running and stationary conditions; Fig. 5g). This did not consistently and significantly improve the response prediction. When comparing the model performance when the running state was included to that of the stimulus-only model, we found that the distribution was largely centered along the diagonal, with a slight asymmetry in favor of the running-dependent model for the better-performing models (Fig. 5h; 28% of responsive neurons had r > 0.5 for stimulus × running state, which was 21% when considering all neurons). This was further corroborated by a simpler model that predicts neural response on the basis of the running speed (rather than a binary condition and without stimulus information) (Supplementary Fig. 7). However, when considering only the 13% of neurons that showed significant modulation of evoked responses (Fig. 5f), the inclusion of running in the categorical model provided a clear predictive advantage (Fig. 5i; the
**Fig. 3** | **Tuning properties reveal functional differences across areas and Cre lines.**

**a.** Pawplot visualization summarizing the median value of a tuning metric across visual areas. Top: each visual area is represented as a circle, with V1 in the center and the higher visual areas surrounding it according to their location on the surface of the cortex. Bottom: each paw-pad (visual area) has two concentric circles. The area of the inner, colored circle, relative to the outer circle, represents the proportion of responsive cells for that layer and area. The color of the inner circle reflects the median value of the metric for the responsive cells. For a metric’s summary plot, four pawplots are shown, one for each layer. Only data from one Cre line are shown for each layer. For each panel, a pawplot is paired with a box plot or a strip plot (for single-cell and population metrics, respectively) showing the full distribution for each Cre line and layer in V1. Data were assigned to cortical layers on the basis of both the Cre line and the imaging depth. Data collected less than 250 µm from the surface were considered to be in layer 2/3, data collected between 250 and 365 µm were considered to be in layer 4, data collected between 375 and 500 µm were considered to be in layer 5, and data collected at 550 µm were considered to be in layer 6. The box shows the quartiles of the data, and the whiskers extend to 1.5 times the interquartile range; points outside this range are shown as outliers. For other cortical areas, see Supplementary Fig. 1.

**b.** Pawplots and box plots summarizing direction selectivity. See Extended Data Fig. 3 for sample sizes.

**c.** Pawplots and box plots summarizing receptive field area. See Extended Data Fig. 5 for sample sizes.

**d.** Pawplots and box plots summarizing preferred temporal frequencies. See Extended Data Fig. 3 for sample sizes.

**e.** Pawplots and box plots summarizing preferred spatial frequencies. See Extended Data Fig. 3 for sample sizes.

**f.** Pawplots and box plots summarizing lifetime sparseness of responses to natural scenes. See Extended Data Fig. 6 for sample sizes.
Fig. 4 | Population correlations have heterogeneous impact on decoding performance. a, Pawplots and strip plots summarizing decoding performance for drifting grating direction using k-nearest-neighbors analysis. Each dot represents the mean fivefold cross-validated decoding performance from a single experiment, with the median performance for a Cre line and layer represented by the bar. See Extended Data Fig. 3 for sample sizes (column ‘expts’). For other cortical areas, see Supplementary Fig. 4. b, Pawplots and strip plots summarizing the population sparseness of responses to natural scenes. See Extended Data Fig. 6 for sample sizes (column ‘expts’). For other cortical areas, see Supplementary Fig. 3. c, Pawplots and strip plots summarizing noise correlations in the responses to drifting gratings. See Extended Data Fig. 3 for sample sizes (column ‘expts’). d, Pawplots and strip plots summarizing the impact of shuffling on decoding performance for drifting grating direction. See Extended Data Fig. 3 for sample sizes (column ‘expts’). Note the diverging colorscale representing both negative and positive values.
mean $r$ for stimulus only was 0.35 and for stimulus $\times$ running was 0.44, whereas the mean $r$ for non-modulated neurons for stimulus only was 0.21 and for stimulus $\times$ running was 0.20).

One of the unique aspects of this dataset is the broad range of stimuli, allowing for a comparison of response characteristics and model predictions across stimuli. Surprisingly, knowing whether a neuron responded to one stimulus type (for example, natural scenes, drifting gratings, etc.) was largely uninformative of whether it responded to another stimulus type. Unlike the examples shown in Fig. 2, which were chosen to highlight responses to all stimuli, most neurons were responsive to only a subset of the stimuli (Fig. 6a). To explore the relationships between neural responses to different types of stimuli, we computed the correlation between the percentage of responsive trials for each stimulus. This comparison removes the threshold of ‘responsiveness’ and examines underlying patterns of activity. We found that most stimulus combinations were weakly correlated (Fig. 6b), demonstrating that knowing that a neuron responds reliably to drifting gratings, for example, carries little to no information about how reliably that neuron responds to one of the natural movies. There was a higher correlation between the reliability of the responses to the natural movie that was repeated across all three sessions (natural movies 1A, 1B and 1C), providing an estimate of the variability introduced by imaging across days and thus a ceiling for the overall correlations across stimuli. Very few of the cross-stimulus correlations approached this ceiling, with the exception of the correlation between static gratings and natural scenes.

We characterized the variability by clustering the reliability, defined by the percentage of significant responses to repeated stimuli. We used a Gaussian mixture model to cluster the 25,958 neurons that were imaged in both sessions A and B (Fig. 1f) and excluded the locally sparse noise stimulus owing to the lack of a comparable definition of reliability. Using neurons imaged in all three sessions did not qualitatively change the results (Supplementary Fig. 8). The clusters are described by the mean percentage of responsive trials for each stimulus for each cluster (Fig. 6c). Note that there was only a weak relationship between the percentage of responsive trials to one stimulus and the percentage of responsive trials to any other stimulus. We grouped the clusters into ‘classes’ by first defining a threshold for responsiveness by identifying the cluster with the lowest mean percentage of responsive trials across stimuli and then setting the threshold equal to the maximum value across stimuli plus 1 s.d. for that cluster. This allowed us to identify each cluster as responsive (or not) to each of the stimuli. Clusters with the same profile (for example, responsive to drifting gratings and natural movies, but not to static gratings or natural scenes) were grouped into one of 16 possible classes.

The clustering was performed 100 times with different initial conditions to evaluate robustness. The optimal number of clusters, evaluated with model comparison, and the class definition threshold were consistent across runs (Supplementary Fig. 8). By far, the largest single class revealed by this analysis was that of neurons that were largely unresponsive to all stimuli, termed ‘none’, which contained 34 ± 2% of the neurons (Fig. 6d). Other large classes included neurons that responded to drifting gratings and natural movies (DG–NM; 14 ± 3%) to natural scenes and natural movies (NS–NM; 14 ± 2%) and to all stimuli (DG–SG–NS–NM; 10 ± 1%).

Notably, we did not observe all 16 possible stimulus response combinations. For instance, very few neurons were classified as responding to one stimulus alone, with the most prominent exception being neurons that responded uniquely to natural movies. Thus, while the pairwise correlations between most stimuli were relatively weak, there was meaningful structure in the patterns of responses. Nevertheless, within each class there remains a great deal of heterogeneity. For example, within the class that responded to all stimuli, there was a cluster in which the neurons responded with roughly equal reliability to all four stimuli (cluster 27 in Fig. 6c) as well as clusters in which the neurons responded reliably to drifting and static gratings and only weakly to natural scenes and natural movies (clusters 25 and 28). This heterogeneity underlies the inability to predict whether a neuron responds to one stimulus given that it responds to another.

Classes were not equally represented in all visual areas (Fig. 6e). The ‘none’ class was larger in the higher visual areas than in V1 and was largest in RL (Fig. 2g). Classes related to moving stimuli, including NM, DG–NM and DG, had relatively flat distributions across the visual areas, excluding RL. The classes responsive to natural stimuli, including NS–NM, DG–NS–NM, SG–NS–NM, and DG–SG–NS–NM, were most numerous in V1 and LM, with lower representation in the other visual areas. This divergence in representation of the motion stimulus response classes and natural stimulus response classes in areas AL, PM, and AM is consistent with the putative dorsal and ventral stream segregation in the visual cortex.

In addition to differential representation across cortical areas, the response classes were also differentially represented among the Cre lines (Fig. 6f). Notably, Sst interneurons in V1 had the fewest neurons labeled as ‘none’ and the most DG–SG–NS–NM neurons. Meanwhile, the plurality of Vip interneurons were in the classes responsive to natural stimuli, specifically natural movies.

---

**Fig. 5 | Neural activity is extremely variable, and this variability is not accounted for by running behavior.** a. Pawplots and box plots summarizing the percentage of responsive trials that had a significant response for each neuron’s preferred drifting grating condition. The responsiveness criterion was that a neuron responded in 25% of the trials; hence, the values in the box plots are bounded at 25%. The box shows the quartiles of the data, and the whiskers extend to 1.5 times the interquartile range; points outside this range are shown as outliers. For box plots for other cortical areas, see Supplementary Fig. 5. See Extended Data Fig. 3 for sample sizes. b. Pawplots and box plots summarizing the coefficient of variation for each neuron’s response to its preferred drifting grating condition. See Extended Data Fig. 3 for sample sizes. c. Two example neurons showing individual trial responses along with the mean tuning curve, where $r$ is the Pearson correlation coefficient between the measured and predicted values. $n = 45$ trials per stimulus condition. d, Pawplots and box plots summarizing the categorical regression, where $r$ is the cross-validated Pearson correlation between model prediction and actual response. Only neurons that were responsive to drifting gratings when using our criterion are included. See Extended Data Fig. 3 for sample sizes. e, Pawplots and box plots summarizing the Pearson correlation of neural activity with running speed. Only neurons in imaging sessions where the running fraction was between 20% and 80% are included (Supplementary Fig. 6). See Extended Data Fig. 8 for sample sizes. For neurons present in multiple sessions that met the running criterion, the mean of their running correlations across these sessions was used here. Note the diverging color scale representing both negative and positive values. f, Density plot of the evoked response to a neuron’s preferred drifting grating condition when the mouse is running (running speed $> 1$ cm s$^{-1}$) compared to when it is stationary (running speed $< 1$ cm s$^{-1}$). Only neurons that were responsive to drifting gratings and had a sufficient number of running and stationary trials for their preferred condition are included; $n = 10,440$. g, Categorical model for two example neurons (as in e) in which the running (blue) and stationary (red) trials have been segregated, where $r$ is the Pearson correlation coefficient between the measured and predicted values. $n = 14$ (left) and 7 (right) trials per condition. h, Density plot of $r$ for the categorical regression for drifting gratings using only the stimulus condition (horizontal axis) and the stimulus condition $\times$ running state (vertical axis). Only neurons that were responsive to drifting gratings and had a sufficient number of running and stationary trials across stimulus conditions are included. $n = 11,799$. i, As in h, except that only neurons that were significantly modulated by running are shown in the density ($n = 2,791$); other neurons are in gray.
Having characterized neurons by their joint reliabilities to multiple stimuli, we next asked to what extent we can predict neural responses, not on a trial-by-trial basis but including the temporal response dynamics, given the stimulus and knowledge of the animal's running condition. We used a model class that remains in widespread use for predicting visual physiological responses and

![Diagram](image-url)

**Legend:**
- **a:** Percent responsive trials
- **b:** Coefficient of variation
- **c:** Categorical model performance (r)
- **d:** Running correlation (r)
- **e:** Run Stationary
- **f:** AU

**Graphs and Tables:**
- **g:** Run Stationary
- **h:** AU
- **i:** AU
that captures both simple and complex cell behaviors. The model structure uses a dense wavelet basis (sufficiently dense to capture spatial and temporal features at the level of mouse visual acuity and temporal response) and computes from this both linear and quadratic features, each of which are summed, along with the binary running trace convolved with a learned temporal filter, and sent through a soft rectification (Fig. 7a). We trained these models on either the collective natural stimuli or the artificial stimuli to predict the extracted event trace. Whereas we found example neurons for which this model worked extremely well (Fig. 7b and Supplementary Figs. 9–11), across the population only 2% of neurons were well fit by this model ($r > 0.5$; 2% of neurons for natural stimuli and 1% of neurons for artificial stimuli; Fig. 7c), with median $r$ values of ~0.2 (natural stimuli). Model performance was slightly higher in V1 than in the higher visual areas and showed little difference across Cre lines. It is also worth noting that there is a great deal of visually responsive activity that is not captured by these models (Supplementary Fig. 5). When comparing the models’ performances across stimulus categories,
we found that the overall distribution of performance for models trained and tested with natural stimuli was higher than for the corresponding models for artificial stimuli (Fig. 7d), in line with previous reports. The running speed of the mouse did not add significant predictive power to the models, as most regression weights were near zero, with the exception of Vip neurons in V1 (Supplementary Fig. 10). Similarly, incorporating pupil area and position had little effect, as was also the case when the quadratic weights were removed at the population level (Supplementary Fig. 10). Well-fit models tended to have sparser weights (Supplementary Fig. 11).
When comparing the model performance for the neurons in each of the classes defined through the clustering analysis, we found that these classes occupied spaces of model performance consistent with their definitions (Fig. 7e–h). The neurons classified as ‘none’ formed a relatively tight cluster and constituted the bulk of the density close to the origin (Fig. 7e). By definition, these neurons had the least response reliability for all stimuli (Fig. 6c) and were likewise the least predictable. Neurons in the NS–NM class showed high model performance for natural stimuli and low performance for artificial stimuli (Fig. 7f). Finally, neurons that reliably responded to all stimuli (DG–SG–NS–NM) showed a broad distribution of model performance, with the highest median performance, equally predicted by both artificial and natural stimuli (Fig. 7g). As running has been shown to influence neural activity in these data independently of visual stimuli (Fig. 5e), one might expect that the ‘none’ class would be composed largely of neurons that are strongly driven by running activity rather than visual stimuli. Instead, we found that the ‘none’ class had one of the smallest median correlations, overall, with the running speed of the mouse, whereas the DG–SG–NS–NM class had the largest correlation (Fig. 7i).

Discussion

Historically, visual physiology has been dominated by single-neuron electrophysiological recordings in which neurons were identified by responding to a test stimulus. The stimulus was then hand-tuned to elicit the strongest reliable response from that neuron, and the experiment proceeded using manipulations around this condition. Such studies discovered many characteristic response properties, namely, that visual responses can be characterized by combinations of linear filters with nonlinearities such as half-wave rectification, squaring, and response normalization or that neurons (in V1 at least) largely cluster into ‘simple’ and ‘complex’ cells. But these studies may have failed to capture the variability of responses, the breadth of features that will elicit a neural response, and the breadth of features that do not elicit a response. This results in systematic bias in the measurement of neurons and a confirmation bias regarding model assumptions. Recently, calcium imaging and denser electrophysiological recordings have enabled large populations of neurons to be recorded simultaneously. Here we scaled calcium imaging, combining standard operating procedures with integrated engineering tools, to address some of the challenges of this difficult technique, as a means to create an unprecedented survey of 59,610 neurons in mouse visual cortex across 243 mice while using a standard and well-studied but diverse set of visual stimuli. This pipeline reduced critical experimental biases by separating quality control of data collection from response characterization. Such a survey is crucial for assessing the successes and shortcomings of contemporary models of the visual cortex.

By using standard noise and grating stimuli, we find many of the standard visual response features, including orientation selectivity, direction selectivity, and spatial receptive fields with opponent on and off subfields (Figs. 2 and 3). On the basis of responses to these stimuli, we observed functional differences in visual responses across cortical areas, layers, and transgenic Cre lines. In a new analysis of overall reliabilities to both artificial and naturalistic stimuli, we find classes of neurons responsive to different constellations of stimuli (Fig. 6). The different classes are largely intermingled and found in all of the cortical areas recorded here, suggesting a largely parallel organization. At the same time, the over-representation of classes responsive to natural movies and motion stimuli in areas AL, PM, and AM relative to the other classes (which are more responsive to spatial stimuli) is consistent with the assignment of these areas to the putative ‘dorsal’ or ‘motion’ stream. The lack of an inverse relationship, wherein spatial information is over-represented relative to motion in a putative ventral stream, likely reflects the fact that we were unable to image the putative ventral areas LI, POR, and P within our cranial window. Area LM has previously been loosely associated with the ventral stream, but with evidence that it is more similar to V1 than other higher-order ventral areas, and our results appear consistent with the latter. Area RL has the largest proportion of neurons in the ‘none’ class, over 85%, in line with the very low percentage of responsive neurons (Fig. 2). It is possible that neurons in this area are specialized for visual features not probed here or that they show a greater degree of multimodality than in the other visual areas, integrating somatosensory and visual features.

One of the unique features of this dataset is that it includes a large number of different transgenic Cre lines for characterization that label specific populations of excitatory and inhibitory neurons. On a coarse scale, excitatory populations behave similarly; however, closer examination reveals distinct functional properties across Cre lines. For instance, neurons expressing Rorb, Scnn1a-Tg3, and Nr5a1, which label distinct layer 4 populations in V1, exhibit distinct spatial and temporal tuning properties (Fig. 3 and Supplementary Figs. 1 and 2), different degrees of running correlation (Fig. 5) and subtle differences in their class distribution (Fig. 6). These differences suggest that there are separate channels of feed-forward information. Similar differences between Fezf2 and Nts1 in V1, which label two distinct populations of corticofugal neurons found in layers 5 and 6, respectively, indicate distinct feedback channels from V1.

The Brain Observatory data also provide a broad survey of visually evoked responses for both Vip and Sst inhibitory Cre lines. Sst neurons are strongly driven by all visual stimuli used here, with the plurality belonging to the DG–SG–NS–NM class (Fig. 5i). Their responses to drifting gratings are particularly robust in that 94% of Sst neurons in V1 are responsive to drifting gratings and respond quite reliably across trials, far more than is seen in the other Cre lines (Extended Data Fig. 3 and Supplementary Fig. 5). Vip neurons, on the other hand, are largely unresponsive to, and even suppressed by, drifting gratings, with only 9% of Vip neurons in V1 labeled responsive. This extreme difference between these two populations is consistent with previous literature examining the size tuning of these interneurons and supports the disinhibitory circuit between them and the excitatory neurons (Fig. 3f). Interestingly, receptive field mapping using locally sparse noise revealed that Vip neurons in V1 have remarkably large receptive field areas, larger than those of both Sst and excitatory neurons (Fig. 3f), in contrast to the smaller summation area for Vip neurons previously measured using windowed drifting gratings. This suggests that Vip neurons respond to smaller features over a larger region of space. Further, both populations show strong running modulation: responses for both correlate more strongly with the mouse’s running speed than is seen for the excitatory populations (Fig. 5e), and a model based solely on the mouse’s running speed does a better job at predicting their activity than for the excitatory populations (Supplementary Fig. 7).

The true test of a model is its ability to predict arbitrary novel responses, in addition to responses from stimuli used for characterization. Even with the inclusion of running, our models predict responses in a minority of neurons (Fig. 7).

Neurons in the DG–SG–NS–NM class were well predicted, with values comparable to those found in primates, for both natural and artificial stimuli (Fig. 7g). On the basis of the way we chose our stimulus parameters, we expected that neurons with a strong ‘classical receptive field’ would be most likely to appear in this class. However, this class constitutes only 10% of the mouse visual cortex (Fig. 6d). Neurons in the NS–NM class show equally high prediction for natural stimuli, but poor prediction for artificial stimuli.
(Fig. 7f). It is possible that these neurons could be classical neurons as well but are tuned for spatial or temporal frequencies that were not included in our stimulus set. As our stimulus parameters were chosen to match previous measurements of mouse acuity, this could suggest that the acuity in mice has been underestimated.

Remarkably, the largest class of neurons was the ‘none’ class, constituting neurons that did not respond reliably to any of the stimuli (34% of neurons). These neurons are the least likely to be described by classical receptive fields, as evidenced by their poor model performance for all stimuli (Fig. 7f). What, then, do these neurons do? It is possible that these neurons are visually driven but are responsive to highly sparse and specific natural features that may arise through hierarchical processing. Indeed, the field has a growing body of evidence that the rodent visual system exhibits sophisticated computations. For instance, neurons as early as V1 show visual responses to complex stimulus patterns. Alternatively, these neurons could be involved in non-visual computation, including behavioral responses such as reward timing and sequence learning, as well as modulation by multimodal sensory stimuli and motor signals. While we found little evidence that these neurons were correlated with a mouse’s running, recent work has found running to be among the least predictive of such motor signals.

We believe that the openly available Allen Brain Observatory provides an important foundational resource for the community. In addition to providing an experimental benchmark, these data serve as a testbed for theories and models. Already, these data have been used by other researchers to develop image processing methods, to examine stimulus encoding and decoding and to test models of cortical computations. Ultimately, we expect that these data will seed as many questions as they answer, fueling others to pursue both new analyses and further experiments to unravel how cortical circuits represent and transform sensory information.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-019-0550-9.

Received: 23 May 2019; Accepted: 28 October 2019; Published online: 16 December 2019

References
1. Hubel, D. & Wiesel, T. Receptive fields of single neurons in the cat’s striate cortex. J. Physiol. 148, 574–591 (1959).
2. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. J. Physiol. 160, 106–154 (1962).
3. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47 (1991).
4. DeCarlo, J. J., Zoccolan, D. & Rust, N. C. How does the brain solve visual object recognition? Neuron 73, 415–434 (2012).
5. Olshausen, B. & Field, D. What is the other 85% of V1 doing? in 23 Problems in Systems Neuroscience (eds van Hemmen, J. & Sejnowski, T.) 182–211 (Oxford University Press, 2006).
6. Masland, R. H. & Martin, P. R. The unsolved mystery of vision. Curr. Biol. 17, R577–R582 (2007).
7. Carandini, M. et al. Do we know what the early visual system does? J. Neurosci. 25, 10577–10597 (2005).
8. Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L. & Reid, R. C. Functional specialization of mouse high spatial cortical areas. Neuron 72, 1025–1039 (2011).
9. Marshall, J. H., Garrett, M. E., Nauhaus, I. & Callaway, E. M. Functional specialization of seven mouse visual cortical areas. Neuron 72, 1040–1054 (2011).
10. Fournier, J., Monier, C., Pananceanu, M. & Frégnac, Y. Adaptation of the simple or complex nature of V1 receptive fields to visual statistics. Nat. Neurosci. 14, 1053–1060 (2011).
43. Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. Nat. Neurosci. 22, 1677–1686 (2019).
44. Petersen, A., Simon, N. & Witten, D. SCALPEL: extracting neurons from calcium imaging data. Ann. Appl. Stat. 12, 2430–2456 (2018).
45. Sheintuch, L. et al. Tracking the same neurons across multiple days in Ca²⁺ imaging data. Cell Rep. 21, 1102–1115 (2017).
46. Ellis, R. J. et al. High-accuracy decoding of complex visual scenes from neuronal calcium responses. Preprint at bioRxiv https://doi.org/10.1101/271296 (2018).
47. Cai, L., Wu, B. & Ji, S. Neuronal activities in the mouse visual cortex predict patterns of sensory stimuli. Neuroinformatics 16, 473–488 (2018).
48. Zylberberg, J. Untuned but not irrelevant: a role for untuned neurons in sensory information coding. Preprint at bioRxiv https://doi.org/10.1101/134379 (2017).
49. Christensen, A. J. & Pillow, J. W. Running reduces firing but improves coding in rodent higher-order visual cortex. Preprint at bioRxiv https://doi.org/10.1101/214007 (2017).
50. Sweeney, Y. & Clopath, C. Population coupling predicts the plasticity of stimulus responses in cortical circuits. Preprint at bioRxiv https://doi.org/10.1101/265041 (2018).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019
Methods

Transgenic mice. All animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at the Allen Institute for Brain Science in compliance with NIH guidelines. Transgenic mouse lines were generated by using conventional and bacterial artificial chromosomal (BAC) transgenic or knock-in strategies as previously described[12,13]. External sources included Cre lines generated as part of the NIH Neuroscience Blueprint Cre Driver Network (http://www.credrivemice.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 (CremERT2) were treated with ~200 μl of tamoxifen solution (0.2 mg g−1 body weight) via oral gavage once per day for five consecutive days to activate Cre recombinase.

We used the transgenic mouse line A939, in which GCaMP6f expression is dependent on the activity of both Cre recombinase and the tetracycline-controlled transactivator protein (TetA)14. A93 mice were first crossed with Camk2a−/− (TA mice), and the double transgenic mice were then crossed with a Cre driver line. For some Cre drivers, we alternatively leveraged the TIGRE2.0 transgenic platform that combines the expression of TetA and GCaMP6f in a single reporter line (Ai488TTL2-GCf6-CD1-1A2). In Cux2-CreERT2Camk2a−/− (TA:A939TTG-CCM6f) mice, expression is regulated by the tamoxifen-inducible Cux2 promoter, the induction of which results in Cre-mediated expression of GCaMP6f predominantly in superficial cortical layers 2, 3 and 4 (ref. 14; Supplementary Fig. 12 and Supplementary Table 1). Both Emx1-ires-Cre-Camk2a−/−;TA:A939 and Scf17a7-IRES-CreCamk2a−/−;TA:A939 are pan-excitatory lines and show expression throughout all cortical layers and subcerebral regions15. Scn1a-Tg3-Cre;Camk2a-tTA;Ai93 mice exhibit GCaMP6f in excitatory neurons26. Vip-IRES-CreAi48 mice exhibit GCaMP6f in Vip-expressing cells by the endogenous promoter/enhancer elements of the vasoactive intestinal polypeptide locus27. Rorb-IRES2-Cre;Cam2a-tTA;Ai93 mice exhibit GCaMP6f in excitatory neurons in cortical layer 4 (dense patches) and layers 5 and 6 (sparse)28. Scnn1a-Tg3-Cre;Camk2a-tTA;Ai93 mice exhibit GCaMP6f in excitatory neurons in cortical layer 4 and in restricted areas within the cortex, in particular in primary sensory cortices. Nr5a1-Cre-Camk2a−/−;TA:A939 mice exhibit GCaMP6f in excitatory neurons in cortical layer 4 (ref. 11). Rbp4-Cre-Camk2a−/−;TA:A939 mice exhibit GCaMP6f in excitatory neurons in cortical layer 5 (ref. 15). Fez2-CreER;Ai48 mice exhibit GCaMP6f in subcerebral projection neurons in layers 5 and 6 (ref. 11). Tlx3-Cre.PL56;Ai48 mice exhibit GCaMP6f primarily restricted to IT corticostriatal in layer 5 (ref. 15). Ntsr1-IRES-Cre, Gn220;Ai93 mice exhibit GCaMP6f in excitatory corticostriatal neurons in cortical layer 6 (ref. 6). We maintained all mice on a reverse 12-h light cycle after surgery and throughout the duration of the experiment and performed all experiments during the dark cycle.

Cross-platform registration. To register data acquired between instruments and repeatedly target and record neurons in brain areas identified with intrinsic imaging, we developed a system for cross-platform registration (Supplementary Fig. 13).

Surgery. Transgenic mice expressing GCaMP6f were weaned and genotyped at approximately postnatal day (P) 21, and surgery was performed between P37 and P63. Surgical eligibility criteria included (1) weight ≥19.5 g (males) or ≥16.7 g (females); (2) normal behavior and activity; and (3) healthy appearance and posture. A preoperative injection of dexamethasone (3.2 mg kg−1, subcutaneously) was administered, an incision was made to remove skin and the exposed skull was leveled with respect to the cranial window.

The disc was covered with a layer of removable foam (Super-Resilient Foam, 8637SK242, McMaster). Data were initially obtained with the mouse eye centered both laterally and vertically on the stimulus monitor and positioned 15 cm from the monitor, with the monitor parallel to the mouse’s body. Later, the monitor was moved to better fill the visual field. The normal distance of the monitor from the eye remained 15 cm, but the monitor 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 1-h imaging sessions at a given FOV during which mice passively observed three different stimuli. One imaging session was performed per mouse per day, for a maximum of 16 sessions per mouse.

On the first day of imaging at a new FOV, 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 an FOV of ~800 μm using epifluorescence and blue light illumination. Once a region was selected, the objective was shielded from stray light coming from the stimulus monitor with 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 and epifluorescence with (2) the pattern of vessels in two-photon imaging and (3) the pattern of cellular labeling 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 (on the basis of a look-up table against depth) and dynamic range while avoiding pixel saturation. The stimulus monitor was clamped in position, and the experiment began. Two-photon movies (512×512 pixels, 30 Hz), eye tracking (30 Hz) and a side-view full-body camera (30 Hz) were recorded. Recording sessions were interrupted and/or failed 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 (>95% of pixel values in a 10-pixel ×10-pixel box in the image was saturated); (3) loss of baseline intensity in excess of 20% caused by bleaching and/or loss of immersion water; and (4) hardware failures causing a loss of data integrity. Immersion water was occasionally supplemented while imaging by using a micropipette taped to the objective (Microfil, MF28676-S1WI) and connected to a 5-mL syringe via an extension tube. At the end of each session, a 2 stack of images (700 μm around the imaging site, 0.1-μm step) was collected to evaluate cortical anatomy and z drift during the course of the experiment. Experiments with z drift above 10 μm over the course of the entire session were excluded. In addition, for each FOV, a full-depth cortical z stack (~700-μm total depth, 5-μm step) was collected to document the imaging site location (Supplementary Figs. 16 and 17).

Detection of epileptic mice for exclusion. Before two-photon imaging, each mouse was screened for the presence of interictal events in two ways. First, on the habituation day on the two-photon rig, we collected a 5-min-long video on the surface of S1 by using the epifluorescence light path of the two-photon rig. For
At least 5 min of spontaneous activity was recorded in each session. ~60-min imaging sessions (Fig. 1f). Session A included drifting gratings and natural movie 1. Session C included locally sparse noise and natural movies 1 and 3. Session B included static gratings, natural scenes and natural movies 1 and 3. The different stimuli were presented in segments of 5–13 min and interleaved. The grating was presented at eight different directions (separated by 45°) and at five temporal frequencies (1, 2, 4, 8 and 15 Hz). Each grating was presented for 2 s, followed by 1 s of mean-luminance gray. Each grating condition was presented 15 times. Trials were randomized, with blank sweeps (that is, 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 spatial frequency (0.04 cycles per degree) and contrast (80%). The grating was presented at eight different directions (separated by 45°) and at five temporal frequencies (1, 2, 4, 8 and 15 Hz). Each grating was presented for 2 s, followed by 1 s of mean-luminance gray. Each grating condition was presented 15 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\textsuperscript{2}*, the van Hateren Natural Image Dataset\textsuperscript{3}, and the McGill Calibrated Colour Image Database\textsuperscript{4}. The images were presented in grayscale and were contrast normalized and resized to 1,174 × 918 pixels. The images were presented for 0.25 s each with no inter-image gray period. Each image was presented ~50 times. Trials were randomized, with blank sweeps presented approximately once every 100 images. Three natural movie clips were used from the opening scene of the movie Touch of Evil (Orson Wells, Universal Pictures, 1958). Natural movies 1 and 2 were both 30-s clips, while natural movie 3 was a 120-s clip. All clips had been contrast-normalized and were presented in grayscale at 30 frames per second. Each movie was presented ten times with no inter-trial gray period. Natural movie 1 was presented in each imaging session. The visually sparse noise stimulus consisted of white and dark spots on a mean-luminance gray background. Each spot was square, with 4.65° on a side. Each frame had ~11 spots on the monitor, with no two spots within 23° of each other, and was presented for 0.25 s. Each of the 16 × 28 spot locations was occupied by white and black spots a variable number of times (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 three-dimensional (3D) image volume of coronal brain images for each specimen. This 3D volume enables spatial registration of each specimen’s cortical laminar structure; and (4) compression of the ipsilateral cortex or the contralateral cortex that resulted in disruption to the visual region, as mapped by ISI. Retinotopic mapping of RL commonly yielded retinotopic centers close to the boundary between RL and somatosensory cortex. Consequently, for some RL experiments, the FOV spanned the boundary between the visual and somatosensory cortex. All RL experiments were reviewed with a semiautomated process (Supplementary Fig. 21), and ROIs that were deemed to lie outside putative visual cortex boundaries (~25%) were excluded from further analysis.

Neuropil subtraction. To correct for contamination of the ROI calcium traces by surrounding neuropil, we modeled the measured fluorescence trace of each cell as $F_{ROI} = F_c + rF_m$, where $F_c$ is the measured fluorescence trace, $F_m$ is the unknown true ROIs by using a model where every ROI had a trace distributed in some spatially heterogeneous, time-dependent fashion $F_c = \sum_i W_i T_{oi} T_{oi}^*$, where $W_i$ is a tensor containing time-dependent weighted masks: $W_i$ measures how much of neuron $i$'s fluorescence is contained in pixel $j$ at time $t$, $T_{oi}$ is the fluorescence trace of neuron $k$ at time $t$—this is what we want to estimate, $F_c$ is the recorded fluorescence in pixel $j$ at time $t$. This model applied to all ROIs before filtering for somas. We filtered out duplicates (defined as two ROIs with >70% overlap) and ROIs that were the union of...
of others (any ROI where the union of any other two ROIs accounted for 70% of its area) before demixing and applied the remaining filtering criteria afterwards. Projecting the movie \( F \) onto the binary masks, \( A \), reduced the dimensionality of the problem from \( 512 \times 512 \) pixels to the number of ROIs.

\[
\sum_i A_i F_{i,j} = \sum_i A_i W_{i,j} T_{i,j}
\]

where \( A_i \) is 1 if pixel \( i \) is in ROI \( A \) and 0 otherwise—these are the masks from segmentation, after filtering. At time point \( t \), this yields the linear regression

\[
A_i(t) = (A_i^T(t)) \bar{T}(t)
\]

where we estimated the weighted masks \( W \) by projection of the recorded fluorescence \( F \) onto the binary masks \( A \). On every frame \( t \), we computed the linear least-squares solution \( \bar{T} \) to extract each ROI's trace value.

It was possible for ROIs to have negative or zero demixed traces \( \bar{T} \). This occurred if there were unions (one ROI composed of two neurons) or duplicates (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 them were removed from the experiment. This led to the loss of ~1% of ROIs (Supplementary Fig. 22).

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

To estimate lambda, which controls the strength of the L0 penalty, we estimated the expected event magnitude, as a function of the number of spikes observed in events smaller than 2 s.d. of the noise, while retaining at least one recovered event.

\[
\text{Lifetime sparseness was computed by using the definition in Vinje and Gallant}^{26}
\]

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

For each stimulus, we computed \( C_{\text{null}} \), the expected correlation between the sample trial-averaged response and the true (unmeasured) mean response. It provides an upper bound on the expected performance of any model that predicts response from the given stimulus trial structure. We followed the computation from Schoppe et al.\(^{76} \)

\[
S_t = \frac{1 - \frac{1}{N} \left( \frac{N}{N-1} \sum_{i=1}^{N} \text{Var}(R_i) \right)}{1 - \frac{1}{N} \left( \frac{N}{N-1} \sum_{i=1}^{N} \text{Var}(R_i) \right)}
\]

where \( N \) is the number of trials and \( R_i \) is the time series of the response on the \( i \)th trial. For \( R_i \), we used the trace of extracted event magnitudes at 30 Hz, smoothed with a Gaussian window of width 0.25 s.

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. To prevent trial-by-trial fluctuations from contaminating our signal correlation estimates, we separated the trials for each stimulus into two subsets and calculated the correlation between the trial-averaged responses with each subset of trials. We averaged the signal correlations over 100 random splits of the trials. 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 ten-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 \( k \)-nearest-neighbors classifiers to decode the visual stimulus identity (for example, the natural scene number, within the natural scenes responses) from the population vector of single-trial responses, by using the correlation distance between response vectors. We report the performance on the heldout data from fivefold cross-validation. On each cross-validation fold, we performed an inner-round of twofold cross-validation to choose the number of neighbors from eight logarithmically spaced options (1, 2, 4, 7, 14 and 27).

Categorical regression model for trial responses to drifting gratings. We fit linear ridge regression models for the trial-averaged responses (events summed during each stimulus presentation). The response for trial \( t \), \( R_t \), is governed by the following equation

\[
R_t = \sum_i w_i'^t s
\]

where \( I_s'^t \) is the characteristic function for stimulus condition \( s \) during trial \( t \), \( I_s'^t \) is equal to 1 when the stimulus condition is equal to \( s \) during trial \( t \) and 0 otherwise, and \( w_i \) is the weight for stimulus condition \( s \) (it gives the response of the neuron to stimulus \( s \)).
We fit two separate models, one for which the stimulus conditions enumerate the different values of the drifting grating (that is, orientations and temporal frequencies, including the blank sweep; 41 total conditions) and another for which each stimulus condition occurred in pairs, one during running and one when the animal was stationary. On each stimulus trial, we classified the locomotion as running or stationary by using a Gaussian mixture model with a Dirichlet process prior for the number of components. Stationary trials were identified by the component with the smallest variance among those with mean speed < 1 cm s⁻¹ (if any existed). We used stimulus conditions with at least five repetitions in each behavioral state and used the same number of trials for each stimulus condition in each behavioral state.

We regressed the summed trial against the combination of stimulus condition and behavioral state (for example, 180 degrees, 4 Hz and running). The regularization weight was chosen by leave-one-out cross-validation on the training data. We also regressed against just the locomotion state, binning the activity and running speed into pseudo-trials of the same length as the drifting grating trials. We measured model performance by the correlation of the predictions and data on holdout trials (fivefold cross-validation).

Regression models for mouse running speed. We performed a polynomial regression of each neuron’s activity against running speed. To do this, we rank-sorted the running speed and binned it into 900-point bins. All speeds between −1 cm s⁻¹ and 1 cm s⁻¹ were labeled stationary. We summed each neuron’s events in the same speed bins to compute the speed tuning. We then fit a polynomial regression for the speed tuning with fivefold cross-validation. On each training fold, we performed an inner twofold cross-validation to select the polynomial degree between 1 and 4. We used ridge regression with leave-one-out cross-validation to compute the speed tuning with fivefold cross-validation. On each stimulus trial, we classified the locomotion state and used the same number of trials for each stimulus condition in each behavioral state.

3D Gabor wavelet model for temporal responses. Each neuron was modeled as a sparse linear combination of basis functions, similar to other approaches. We used a pyramid of 3D Gabor wavelet filters that tiled the stimulus at multiple scales, directions, and temporal frequencies (Fig. 6a). The filters are defined by

\[ f(x, y, t; \lambda, \theta, \psi, \sigma, \tau) = \exp\left(-\sigma^2 + \psi^2 + \tau^2\right) \exp\left(i \left(\frac{2\pi x^2}{\sigma^2} + \psi t\right)\right) \]

where \( z = x \cos \theta + y \sin \theta \)
\( y' = -x \sin \theta + y \cos \theta \)

\( \lambda \) controls spatial frequency, \( \theta \) orientation, \( \psi \) temporal frequency, \( \sigma \) the Gaussian envelope, and \( \tau \) the Gaussian envelope in time. This linear model forms a reasonably tight frame. The parameters that generate the set of filters were adapted and scaled to the tuning properties of mouse visual cortex. We estimated weights for ten time lags for each basis function to enable fitting of the temporal kernel. The weighted sum of the basis functions was passed to a polynomial regression.

The model is technically a generalized linear model (where the linear model is built by considering linear combinations of the features \( H(t) \) and \( H^2(t) \), along with a temporal filter for the running signal of the animal, \( r(t) \), with a parameterized soft-plus output. The weights \( w \) were fit to the data by using threshold gradient descent and the Poisson negative log-likelihood cost function, with rate

\[ \tilde{R}(t) = \log \frac{1}{1 + \exp \left( \sum \lambda_y H_y(t) \right)} + H^2(t) \left(1 - r(t) \right)w_0^2(t) + r(t) \left(1 - w(t) \right) \]

This model, with a quadratic dependence on the stimulus \( H^2(t) \), is akin to a regularized STA/STC analysis, adapted to fit the full spatiotemporal receptive field by using stimuli from the data.

We estimated a sparse combination of basis functions for each neuron by using a variant of threshold gradient descent. In threshold gradient descent, only basis functions whose gradient exceeded the threshold had its weight added to the ‘active set’, which was maintained over the optimization, and then all weights in the active set were updated, preventing oscillations. Third, we used an adaptive step size. The step size increased by a factor of 1.2 at each iteration, and the step size decreased by a factor of 0.5 if generalization worsened.

We used a nested sixfold cross-validation framework. We split the data into six sets each containing many 50-sample-long continuous blocks from throughout the dataset. 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 set functioning as the stopping set. The five models were averaged together to make predictions on the test set. Reported model performance is the average on the test set across the six folds. Separate models were fit for the natural stimuli and the artificial stimuli. The weights for these models were sparse, and for all models fewer than 20% of the basis functions had non-zero weight values (Supplementary Fig. 1). The number of parameters for the model was 317,451.

Pupil position and area were measured (Supplementary Fig. 24), and these were incorporated into some of the models. These corrections had little effect on model performance (Supplementary Fig. 10).

We show examples of this model on four neurons in Supplementary Fig. 9.

Clustering of reliabilities. We performed a clustering analysis using the reliabilities by stimulus for each cell (defined as the percentage of responsive trials to the cell’s preferred stimulus condition). We did not include locally sparse noise in this analysis. We combined the reliabilities for natural movies by taking the maximum reliability across the different stimulus conditions. We performed a Gaussian mixture model clustering on these reliabilities for cluster number ranging from 1 to 4, by using the Bayes Information Criterion on data with fourfold cross-validation to select the optimal number of clusters. Once the optimal model was selected, we defined a threshold for responsiveness by selecting the cluster with the lowest mean reliability over all stimuli. We set the threshold to be the maximum reliability plus 1 s.d. over the reliabilities for this cluster. By using this threshold, we identified each cluster according to its profile of responsiveness (that is, whether it responded to drifting gratings, etc.), defining these profiles as ‘classes.’ For each cell, we predicted cluster membership by using the optimal model and then the class membership by using the threshold. We repeated this process 100 times to estimate the robustness of the clustering and derive uncertainties for the number of cells belonging to each class.

Statistics. No statistical methods were used to predetermine sample sizes per location, but our sample sizes are similar to those reported in previous publications.

Data collection and analyses were not performed with blinding to the conditions of the experiments, as there was a single experimental condition for all acquired data. Within each transgenic Cre line, mice were randomly assigned to data collection to sample different areas and imaging depths. Stimulus conditions for gratings and natural scenes were presented in a randomized order within each epoch, as described. No other randomization was used as there were fixed experimental conditions for all other aspects of the dataset. Additional research design information can be found in the Nature Research Reporting Summary accompanying this study.
from these tests were then corrected with the Sidak method to account for multiple comparisons. If the P value for any patch on the stimulus monitor was significant (P < 0.05) after multiple-comparison correction, the neuron was considered to have a receptive field.

Test for inclusion of locally sparse noise spots in a receptive field. The receptive field was computed by using an event-triggered average. Because more than one stimulus spot was present during a given trial, it is not possible to infer the stimulus-response relationship between spot locations and responses on a per-trial basis. Therefore, a statistically significant co-occurrence of spot presentation and responses across trials defined the inclusion criteria for membership of a stimulus spot 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 spots. A P value 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 receptive trials. This per-spot null distribution was estimated by shuffling the identity of the responsive trials (n = 10,000 shuffles). Statistical outliers were identified by computing a P value for each spot relative to its null distribution. These P values were corrected for false discoveries by using the Sidak method to identify receptive field membership.

Comparison of single-cell response metric distributions. To compare the distributions of single-cell response metrics across areas, layers and Cre lines, we used a Kolmogorov–Smirnov test with a Bonferroni correction for the number of comparisons, defined as the number of other distributions to which we were comparing, for example, for area-wise comparison of the Cux2 line, there were six areas in total and thus five comparisons for each area (first row of Supplementary Fig. 2). The Kolmogorov–Smirnov test was chosen as it does not assume a normal distribution nor equal variance.

To compare the distributions of 265

The Kolmogorov–Smirnov test with a Bonferroni correction for the number of comparisons correction and thresholded at

values were corrected for false discoveries by using the Šidák multiple-comparison correction and thresholded at

represents project administration. S.E.J.d.V., J.A.L. and M.A.B. wrote the paper with input from P.A.G., G.K.O., M.O., N.C., P.L., D.M., R.C.R., M.G. and C.K. 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).

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

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

Jewell, S. & Witten, D. Exact spike train invariance via ε, optimization. Ann. Appl. Stat. 12, 2457–2482 (2018).

Oliphant, T. E. A Guide to NumPy (Trelgol Publishing, 2006).

Jones, E., Oliphant, T., Peterson, P. et al. SciPy: Open source scientific tools for Python. http://www.scipy.org/ (2001).

McKinney, W. Data structures for statistical computing in Python. Proceedings of the 9th Python in Science Conference (eds van der Walt, S. & Millman, J.) 51–56 (2010).

Hunter, J. D. Matplotlib: a 2D graphics environment. Comput. Sci. Eng. https://doi.org/10.1109/MCSE.2007.55 (2017).

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).

Kay, K. N., Naselaris, T., Prenger, R. J. & Gallant, J. L. Identifying natural images from human brain activity. Nature 452, 352–355 (2008).

Nishimoto, S. et al. Reconstructing visual experiences from brain activity evoked by natural movies. Curr. Biol. 21, 1641–1646 (2011).

Willmore, B. D. B., Prenger, R. J. & Gallant, J. L. Neural representation of natural images in visual area V2. J. Neurosci. 30, 2102–2114 (2010).

Friedman, J. H. & Popescu, B. E. Predictive learning via rule ensembles. Ann. Appl. Stat. 2, 916–954 (2008).

Riddermeer, M. & Braun, H. BPROP—a fast adaptive learning algorithm. Proceedings of the International Symposium on Computer and Information Science VII (1992).

Teeters, J. L. et al. Neurodata Without Borders: creating a common data format for neurophysiology. Neuron 88, 629–634 (2015).

Author contributions

S.E.J.d.V., M.A.B., K.R., M.G., T.K., S.M., S.O., J.W., C.H., D.L., J.A., A.W. and C.K. conceived of and designed the experiment. J.A.L., T.K., P.H., A.L., C.S., D.S., H.Z., C.D., L.N., A.B., J.W.P., R.C.R., M.G., T.K., S.M., S.O., J.W., H.Z., C.D., L.N., A.B., J.W.P., R.C.R. built and maintained the hardware. S.E.J.d.V., J.A.L., M.A.B., G.K.O., D.F., N.C., L.K., W.W., D.W., R.V., C.B., B.B., T.D., J.G., T.G., S.J., N.K., C.L., F. Lee, F. Long, R.H., L.H., U.K., J.L., J.D.L., R.L., E.L., L.L., J.L., K.M., T.N., M.R., S.S., C.W. and A.W. provided project administration. S.E.J.d.V., J.A.L. and M.A.B. wrote the paper with input from P.A.G., G.K.O., M.O., N.C., P.L., D.M., R.C.R., M.G. and C.K.

References

51. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140 (2010).
52. Madisen, L. et al. Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuro Image 85, 942–958 (2015).
53. Dagle, T. L. et al. A suite of transgenic driver and reporter mouse lines with enhanced brain-cell-type targeting and functionality. Cell 174, 465–480 (2018).
54. Franco, S. J. et al. Fate-restricted neural progenitors in the mammalian cerebral cortex. Science 337, 746–749 (2012).
55. Harris, J. A. et al. Anatomical characterization of Cre driver mice for neural circuit mapping and manipulation. Front. Neural Circuits 8, 1–16 (2014).
56. Gorski, J. A. et al. Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing line. J. Neurosci. 22, 6309–6314 (2002).
57. Taniguchi, H. et al. A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron 71, 995–1013 (2011).
58. Dhillon, H. et al. Leptin directly activates SF1 neurons in the VMH, and this action by leptin is required for normal body-weight homeostasis. Neuron 49, 191–203 (2006).
59. Gerfen, C. R., Paletzki, R. & Heintz, N. GENSAT BAC Cre recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. Neuron 80, 1368–1383 (2015).
60. Guo, C. et al. Fezf2 expression identifies a multipotent progenitor for neocortical projection neurons, astrocytes, and oligodendrocytes. Neuron 80, 1167–1174 (2013).
Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41593-019-0550-9.

Supplementary information is available for this paper at https://doi.org/10.1038/s41593-019-0550-9.

Correspondence and requests for materials should be addressed to S.E.J.d., J.A.L. or M.A.B.

Reprints and permissions information is available at www.nature.com/reprints.
Extended Data Fig. 1 | Spontaneous and evoked event magnitude. **a,** Pawplot and box plots summarizing the mean event magnitude for neurons during the 5 minute spontaneous activity (mean luminance gray) stimulus. For a description of the visualization see Fig. 3. The box shows the quartiles of the data, and the whiskers extend to 1.5 times the interquartile range. Points outside this range are shown as outliers. See Extended Data Figure 3 for sample sizes. **b,** Pawplot and box plots summarizing the maximum evoked event magnitude for neurons’ responses to drifting gratings. See Extended Data Figure 3 for sample sizes.
Extended Data Fig. 2 | Response visualizations. Conventional tuning curves for drifting grating responses for one neuron. a, Direction tuning plotted at the preferred temporal frequency (4 Hz) (mean ± sem across 15 trials). Dotted line represents the mean response to the blank sweep. b, Temporal frequency tuning plotted at the preferred grating direction (270°) (mean ± sem). 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 neuron. g, Orientation tuning plotted at the preferred spatial frequency (0.04 cpd) for each of the four phases. (mean ± sem across 50 trials) Dotted line represents the mean response to the blank sweep. h, Spatial frequency tuning plotted at the preferred orientation (90°) for each of the four phases (mean ± sem). 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 for one neuron. m, Responses to each image presented (mean ± sem across 50 trials). Dotted line represents the mean response to the blank sweep. n, All trials of the image which elicited the largest mean response, the 250ms 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. o, All trials are clustered, with the strongest response in the center and weaker responses on the outside. Clusters are placed on a “Corona plot”. 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 for one neuron. p, Responses of one neuron’s response to each of 10 trials of the natural movie. q, 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.
### Extended Data Fig. 3 | Responsiveness to drifting gratings.

**a.** Table summarizing the numbers of experiments (expts) and neurons imaged for each Cre line, layer, area combination in response to drifting grating stimulus and the number, and percent, of neurons that were responsive to the drifting grating stimulus.

**b.** Strip plots of the percent of neurons responsive to the drifting grating stimulus for each experiment.

| Cre | V1 | LM | AL | PM | AM | RL |
|-----|----|----|----|----|----|----|
|     | expts | total | responsive | percent | expts | total | responsive | percent | expts | total | responsive | percent | expts | total | responsive | percent | expts | total | responsive | percent | expts | total | responsive | percent |
| Enx1 - layer 2/3 | 3 | 628 | 341 | 54% | 5 | 769 | 493 | 59% | 2 | 328 | 189 | 58% | 2 | 293 | 104 | 35% | 1 | 109 | 49 | 46% | 4 | 446 | 100 | 22% |
| Enx1 - layer 4 | 3 | 625 | 306 | 49% | 3 | 677 | 359 | 53% | 2 | 581 | 413 | 70% | 1 | 188 | 82 | 44% | 1 | 83 | 22 | 27% | 4 | 623 | 209 | 25% |
| Enx1 - layer 6 | 4 | 740 | 414 | 56% | 3 | 318 | 187 | 59% | 1 | 75 | 53 | 71% | 1 | 70 | 50 | 57% | 1 | 130 | 22 | 17% | 4 | 623 | 209 | 25% |
| Scl17a7 - layer 2/3 | 6 | 955 | 536 | 56% | 6 | 865 | 342 | 40% | 1 | 45 | 19 | 42% | 5 | 342 | 105 | 30% | 2 | 138 | 51 | 37% | 1 | 45 | 17 | 36% |
| Scl17a7 - layer 4 | 6 | 1538 | 963 | 62% | 5 | 1177 | 987 | 85% | 1 | 182 | 104 | 57% | 4 | 363 | 186 | 43% | 1 | 52 | 27 | 40% | 6 | 339 | 197 | 58% |
| Scl17a7 - layer 5 | 6 | 888 | 598 | 67% | 6 | 540 | 209 | 39% | 1 | 12 | 5 | 42% | 6 | 170 | 75 | 44% | 1 | 50 | 25 | 50% | 6 | 339 | 197 | 58% |
| Cux2 - layer 2/3 | 8 | 1919 | 792 | 41% | 8 | 794 | 399 | 50% | 6 | 799 | 371 | 46% | 9 | 949 | 404 | 43% | 1 | 362 | 116 | 41% | 9 | 376 | 107 | 28% |
| Cux2 - layer 4 | 8 | 1939 | 1181 | 61% | 5 | 1228 | 576 | 48% | 7 | 1420 | 815 | 58% | 4 | 676 | 256 | 38% | 6 | 678 | 339 | 40% | 6 | 705 | 194 | 27% |
| Robb | 8 | 1817 | 1013 | 56% | 9 | 767 | 391 | 51% | 6 | 794 | 399 | 49% | 7 | 511 | 250 | 49% | 8 | 516 | 252 | 49% | 5 | 743 | 188 | 25% |
| Sont1a | 9 | 1290 | 317 | 25% | 6 | 461 | 146 | 33% | 6 | 256 | 67 | 26% | 6 | 178 | 73 | 41% | 7 | 203 | 65 | 32% | 6 | 110 | 39 | 30% | 6 | 670 | 249 | 36% |
| Nr3a1 | 6 | 405 | 242 | 60% | 7 | 406 | 229 | 57% | 5 | 397 | 187 | 47% | 6 | 337 | 196 | 58% | 5 | 397 | 187 | 47% | 6 | 337 | 196 | 58% |
| Rbp4 | 7 | 320 | 210 | 66% | 7 | 333 | 189 | 56% | 6 | 287 | 105 | 37% | 6 | 379 | 196 | 52% | 8 | 244 | 136 | 55% | 4 | 64 | 22 | 34% |
| Fzd9 | 4 | 285 | 211 | 74% | 5 | 500 | 408 | 84% | 3 | 684 | 408 | 59% | 3 | 684 | 408 | 59% | 3 | 684 | 408 | 59% | 3 | 684 | 408 | 59% |
| Tn3 | 6 | 840 | 401 | 48% | 3 | 684 | 408 | 59% | 5 | 397 | 187 | 47% | 6 | 337 | 196 | 58% | 5 | 397 | 187 | 47% | 6 | 337 | 196 | 58% |
| Nbr1 | 6 | 405 | 242 | 60% | 7 | 406 | 229 | 57% | 5 | 397 | 187 | 47% | 6 | 337 | 196 | 58% | 5 | 397 | 187 | 47% | 6 | 337 | 196 | 58% |
| Slc17a7 - layer 2/3 | 9 | 101 | 94 | 90% | 8 | 121 | 74 | 61% | 6 | 70 | 39 | 55% | 8 | 81 | 62 | 77% | 1 | 15 | 19 | 100% | 8 | 81 | 62 | 77% | 1 | 15 | 19 | 100% |
| Slc17a7 - layer 4 | 8 | 77 | 73 | 96% | 7 | 82 | 56 | 68% | 8 | 70 | 39 | 55% | 8 | 81 | 62 | 77% | 1 | 15 | 19 | 100% | 8 | 81 | 62 | 77% | 1 | 15 | 19 | 100% |
| Slc17a7 - layer 5 | 9 | 118 | 95 | 81% | 8 | 137 | 13 | 9% | 7 | 108 | 48 | 44% | 8 | 137 | 13 | 9% | 7 | 108 | 48 | 44% | 8 | 137 | 13 | 9% | 7 | 108 | 48 | 44% |
| Vip - layer 2/3 | 9 | 83 | 12 | 14% | 9 | 105 | 26 | 25% | 9 | 130 | 62 | 50% | 9 | 130 | 62 | 50% | 9 | 130 | 62 | 50% | 9 | 130 | 62 | 50% |
| Vip - layer 4 | 9 | 83 | 12 | 14% | 9 | 105 | 26 | 25% | 9 | 130 | 62 | 50% | 9 | 130 | 62 | 50% | 9 | 130 | 62 | 50% | 9 | 130 | 62 | 50% |

**Table summarizing the numbers of experiments (expts) and neurons imaged for each Cre line, layer, area combination in response to drifting grating stimulus and the number, and percent, of neurons that were responsive to the drifting grating stimulus.**

**Strip plots of the percent of neurons responsive to the drifting grating stimulus for each experiment.**
### Extended Data Fig. 4 | Responsiveness to static grating

#### a. Table summarizing the numbers of experiments and neurons imaged for each Cre line, layer, area combination in response to static grating stimulus and the number, and percent, of neurons that were responsive to the static grating stimulus.

| Cre Line | Layer 2/3 | Layer 4 | Layer 5 | Layer 6 |
|----------|-----------|---------|---------|---------|
|           | Neurons  | Neurons | Neurons | Neurons |
| Emx1      | 3        | 379     | 123     | 123     |
| Slic17a7  | 3        | 431     | 165     | 165     |
| Cux2      | 3        | 584     | 244     | 244     |
| Vip       | 3        | 388     | 497     | 497     |
| Emx1      | 4        | 388     | 497     | 497     |
| Slic17a7  | 4        | 388     | 497     | 497     |
| Cux2      | 4        | 388     | 497     | 497     |
| Vip       | 4        | 388     | 497     | 497     |

#### b. Strip plots of the percent of neurons responsive to the static grating stimulus for each experiment.
### Extended Data Fig. 5 | Responsiveness to locally sparse noise

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

| Cre line | Layer | Area | Expts | Total | Responsive | Percent | Expts | Total | Responsive | Percent | Expts | Total | Responsive | Percent | Expts | Total | Responsive | Percent | Expts | Total | Responsive | Percent | Expts | Total | Responsive | Percent |
|----------|-------|------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|
| Emx1 - layer 2/3 | 3 | 949 | 1995 | 30 | 954 | 518 | 79% | 2 | 316 | 257 | 81% | 2 | 270 | 60 | 22% | 1 | 78 | 19 | 24% | 4 | 500 | 39 | 8% |
| Emx1 - layer 4 | 3 | 705 | 661 | 79% | 3 | 684 | 407 | 79% | 2 | 527 | 402 | 76% | 1 | 206 | 26 | 17% | 1 | 117 | 13 | 15% | 4 | 661 | 103 | 12% |
| Emx1 - layer 5 | 4 | 640 | 488 | 79% | 3 | 648 | 407 | 79% | 2 | 573 | 450 | 80% | 1 | 75 | 17 | 22% | 3 | 31 | 2 | 8% | 1 | 120 | 1 | 1% |
| Slic17a7 - layer 2/3 | 6 | 925 | 866 | 95% | 6 | 666 | 489 | 73% | 1 | 35 | 15 | 30% | 6 | 344 | 109 | 32% | 2 | 149 | 34 | 23% | 1 | 34 | 6 | 26% |
| Slic17a7 - layer 4 | 6 | 1394 | 1150 | 82% | 5 | 1380 | 991 | 74% | 1 | 220 | 81 | 35% | 4 | 258 | 43 | 14% | 1 | 34 | 2 | 8% |
| Slic17a7 - layer 5 | 5 | 869 | 860 | 95% | 5 | 284 | 230 | 79% | 6 | 578 | 135 | 23% | 1 | 34 | 2 | 6% |
| Cux2 - layer 2/3 | 8 | 1363 | 427 | 31% | 8 | 617 | 301 | 37% | 6 | 742 | 93 | 13% | 9 | 762 | 52 | 7% | 5 | 354 | 59 | 17% | 6 | 371 | 32 | 9% |
| Cux2 - layer 4 | 8 | 1775 | 884 | 50% | 5 | 651 | 239 | 34% | 7 | 1144 | 413 | 36% | 4 | 537 | 18 | 3% | 6 | 517 | 87 | 17% | 6 | 734 | 52 | 7% |
| Rorb | 8 | 1414 | 967 | 48% | 6 | 785 | 798 | 38% | 6 | 753 | 163 | 22% | 7 | 404 | 48 | 10% | 8 | 434 | 16 | 4% | 5 | 829 | 52 | 7% |
| Snc1a | 8 | 1178 | 921 | 27% | 8 | 778 | 161 | 21% | 6 | 287 | 58 | 22% | 6 | 136 | 31 | 20% | 7 | 214 | 38 | 18% | 7 | 615 | 10 | 5% | 6 | 611 | 22 | 7% |
| Ntsr1 | 8 | 371 | 671 | 19% | 7 | 365 | 96 | 24% | 5 | 363 | 20 | 6% | 5 | 363 | 20 | 6% | 5 | 363 | 20 | 6% | 5 | 363 | 20 | 6% |
| Sat - layer 4 | 9 | 157 | 69 | 46% | 8 | 104 | 48 | 46% | 1 | 14 | 0 | 0% | 6 | 63 | 5 | 5% | 1 | 25 | 0 | 0% | 1 | 25 | 0 | 0% |
| Sat - layer 3 | 6 | 37 | 40 | 79% | 7 | 71 | 28 | 20% | 6 | 63 | 9 | 14% | 1 | 7 | 0 | 0% |
| Vip - layer 3/2 | 9 | 120 | 82 | 68% | 8 | 118 | 65 | 59% | 7 | 105 | 8 | 5% | 7 | 105 | 8 | 5% |
| Vip - layer 4 | 8 | 120 | 64 | 52% | 9 | 84 | 60 | 66% | 9 | 121 | 10 | 8% | 9 | 121 | 10 | 8% |
| Total | 123 | 14504 | 7754 | 53% | 191 | 8927 | 4981 | 55% | 41 | 4838 | 1993 | 34% | 87 | 5270 | 833 | 12% | 30 | 2509 | 258 | 11% | 40 | 4281 | 325 | 8% |
### Extended Data Fig. 6 | Responsiveness to natural scenes.

**a**, Table summarizing the numbers of experiments (expts) and neurons imaged for each Cre line, layer, area combination in response to locally sparse noise stimulus and the number, and percent, of neurons that were responsive to the locally sparse noise stimulus.

**b**, Strip plots of the percent of neurons responsive to the locally sparse noise stimulus for each experiment.

| Cre | expts | total | responsive | percent | expts | total | responsive | percent | expts | total | responsive | percent | expts | total | responsive | percent | expts | total | responsive | percent |
|-----|-------|-------|------------|---------|-------|-------|------------|---------|-------|-------|------------|---------|-------|-------|------------|---------|-------|-------|------------|---------|
| Emx1 - layer 2/3 | 3 | 579 | 346 | 60% | 5 | 721 | 394 | 55% | 2 | 340 | 191 | 56% | 2 | 251 | 123 | 49% | 1 | 88 | 28 | 32% | 4 | 465 | 33 | 7% |
| Emx1 - layer 4 | 3 | 431 | 264 | 60% | 5 | 634 | 307 | 49% | 2 | 943 | 321 | 43% | 1 | 188 | 59 | 31% | 1 | 134 | 29 | 21% | 4 | 750 | 50 | 7% |
| Emx1 - layer 6 | 4 | 584 | 420 | 65% | 6 | 645 | 349 | 54% | 1 | 38 | 22 | 58% | 5 | 248 | 181 | 48% | 2 | 137 | 90 | 67% | 1 | 31 | 14 | 42% |
| Sst17a7 - layer 2/3 | 6 | 966 | 556 | 60% | 6 | 648 | 349 | 54% | 1 | 38 | 22 | 58% | 5 | 248 | 181 | 48% | 2 | 137 | 90 | 67% | 1 | 31 | 14 | 42% |
| Sst17a7 - layer 4 | 6 | 1928 | 860 | 65% | 5 | 1163 | 443 | 54% | 1 | 242 | 71 | 29% | 4 | 288 | 92 | 32% | 1 | 46 | 20 | 42% |
| Sst17a7 - layer 5 | 5 | 587 | 661 | 74% | 5 | 352 | 240 | 68% | 6 | 528 | 280 | 53% | 1 | 48 | 29 | 42% |
| Cux2 - layer 2/3 | 8 | 1931 | 1196 | 61% | 6 | 775 | 432 | 56% | 6 | 763 | 399 | 42% | 9 | 812 | 393 | 49% | 5 | 305 | 171 | 44% | 6 | 383 | 51 | 14% |
| Cux2 - layer 4 | 8 | 1733 | 1090 | 63% | 5 | 1098 | 570 | 44% | 7 | 1158 | 511 | 44% | 4 | 532 | 182 | 35% | 6 | 534 | 155 | 29% | 6 | 673 | 30 | 4% |
| Rorb | 8 | 1404 | 646 | 60% | 6 | 755 | 443 | 59% | 6 | 761 | 360 | 48% | 7 | 431 | 146 | 34% | 8 | 388 | 86 | 22% | 5 | 177 | 42 | 5% |
| Sont1a | 8 | 1114 | 849 | 59% | 6 | 592 | 219 | 36% | 6 | 592 | 219 | 36% | 7 | 228 | 80 | 35% | 6 | 123 | 28 | 23% | 6 | 655 | 34 | 4% |
| Nr1a1 | 6 | 396 | 230 | 59% | 6 | 240 | 91 | 38% | 6 | 110 | 42 | 37% | 7 | 228 | 80 | 35% | 6 | 123 | 28 | 23% | 6 | 655 | 34 | 4% |
| Rbp4 | 7 | 202 | 161 | 69% | 7 | 206 | 206 | 98% | 6 | 235 | 169 | 60% | 6 | 202 | 123 | 35% | 8 | 212 | 94 | 44% | 4 | 51 | 12 | 29% |
| Fezf2 | 4 | 291 | 252 | 89% | 5 | 992 | 474 | 48% | 6 | 829 | 460 | 54% | 3 | 612 | 490 | 30% | 0 | 249 | 97 | 28% |
| Tpx3 | 6 | 629 | 660 | 81% | 5 | 612 | 490 | 30% | 0 | 249 | 97 | 28% |

Extended Data Fig. 6 | Responsiveness to natural scenes. **a**, Table summarizing the numbers of experiments (expts) and neurons imaged for each Cre line, layer, area combination in response to locally sparse noise stimulus and the number, and percent, of neurons that were responsive to the locally sparse noise stimulus. **b**, Strip plots of the percent of neurons responsive to the locally sparse noise stimulus for each experiment.
Extended Data Fig. 7 | Responsiveness to natural movies. a, Table summarizing the numbers of experiments (expts) and neurons imaged for each Cre line, layer, area combination in response to any of the natural movie stimuli and the number, and percent, of neurons that were responsive to the natural movie stimuli. b, Strip plots of the percent of neurons responsive to the natural movie stimuli for each experiment.

| Cre       | expts | V1    | total | responsive | percent | LM    | total | responsive | percent | AL    | total | responsive | percent | PM    | total | responsive | percent | AM    | total | responsive | percent | RL    | total | responsive | percent |
|-----------|-------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|-------|-------|-----------|---------|
| Emx1 -    | 3     | 933   | 512   | 35%       | 512     | 5     | 1122  | 623      | 56%     | 2     | 512   | 307      | 60%     | 2     | 422   | 112      | 27%     | 1     | 107   | 74       | 49%     | 4     | 712   | 99       | 14%     |
| layer 2/3 |       |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |
| Emx1 -    | 3     | 1043  | 700   | 67%       | 700     | 3     | 979   | 479      | 48%     | 2     | 838   | 484      | 58%     | 1     | 308   | 73       | 24%     | 1     | 195   | 29       | 15%     | 4     | 1223  | 129      | 11%     |
| layer 4   |       |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |
| Emx1 -    | 3     | 1089  | 535   | 49%       | 535     | 3     | 979   | 479      | 48%     | 2     | 838   | 484      | 58%     | 1     | 308   | 73       | 24%     | 1     | 195   | 29       | 15%     | 4     | 1223  | 129      | 11%     |
| layer 5   |       |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |
| Scl17a7   | 1      | 1433  | 301   | 65%       | 1433    | 6     | 1115  | 482      | 43%     | 1     | 64    | 31       | 48%     | 5     | 561   | 222      | 38%     | 2     | 235   | 72       | 31%     | 1     | 40    | 8        | 12%     |
| layer 3/5 |       |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |
| Scl17a7   | 1      | 2122  | 1499  | 69%       | 2122    | 6     | 1645  | 1031     | 63%     | 1     | 310   | 59       | 32%     | 4     | 581   | 123      | 22%     | 1     | 72    | 24       | 33%     |
| layer 4   |       |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |
| Scl17a    | 1      | 1285  | 640   | 73%       | 1285    | 5     | 470   | 239      | 71%     | 6     | 1228  | 629      | 49%     | 9     | 1403  | 996      | 38%     | 8     | 819   | 237      | 38%     | 9     | 558   | 144      | 26%     |
| layer 5   |       |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |
| Cu22      | 1      | 2862  | 1561  | 62%       | 2862    | 5     | 1564  | 872      | 55%     | 7     | 1077  | 929      | 49%     | 4     | 928   | 315      | 34%     | 6     | 907   | 311      | 31%     | 6     | 1020  | 144      | 14%     |
| layer 6   |       |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |       |       |           |         |
| Rorb      | 1      | 2218  | 1383  | 62%       | 2218    | 9     | 1191  | 648      | 54%     | 6     | 1242  | 543      | 44%     | 7     | 794   | 300      | 40%     | 8     | 739   | 313      | 43%     | 5     | 1120  | 168      | 15%     |
| Snta1     | 1      | 1873  | 1358  | 56%       | 1873    | 6     | 1191  | 648      | 54%     | 6     | 1242  | 543      | 44%     | 7     | 794   | 300      | 40%     | 8     | 739   | 313      | 43%     | 5     | 1120  | 168      | 15%     |
| Nrta1     | 1      | 578   | 424   | 73%       | 578     | 6     | 421   | 213      | 51%     | 6     | 220   | 116      | 53%     | 7     | 311   | 129      | 38%     | 6     | 171   | 77       | 49%     | 6     | 1394  | 268      | 20%     |
| Rbp4      | 1      | 459   | 304   | 66%       | 459     | 7     | 465   | 293      | 62%     | 6     | 441   | 207      | 61%     | 5     | 509   | 341      | 67%     | 6     | 355   | 170      | 49%     | 4     | 93    | 36       | 41%     |
| Ptz2      | 1      | 407   | 292   | 72%       | 407     | 3     | 981   | 842      | 55%     | 41    | 7101  | 3660     | 54%     | 87   | 7365  | 252      | 41%     | 36   | 3569  | 1388     | 39%     | 40   | 9486  | 1065     | 17%     |

Extended Data Fig. 7 | Responsiveness to natural movies. a, Table summarizing the numbers of experiments (expts) and neurons imaged for each Cre line, layer, area combination in response to any of the natural movie stimuli and the number, and percent, of neurons that were responsive to the natural movie stimuli. b, Strip plots of the percent of neurons responsive to the natural movie stimuli for each experiment.
## Running correlations

| Cre            | V1 | LM | AL | PM | AM | RL |
|----------------|----|----|----|----|----|----|
|                | expts | neurons | expts | neurons | expts | neurons | expts | neurons | expts | neurons | expts | neurons |
| Emx1 - layer 2/3 | 2  | 254 | 1  | 254 | 1  | 143 | 2  | 327 |
| Emx1 - layer 4  | 1  | 148 | 1  | 254 | 1  | 117 | 2  | 322 |
| Emx1 - layer 5  | 1  | 334 | 2  | 296 | 1  | 53  | 1  | 162 |
| Slic17a7 - layer 2/3 | 3  | 656 | 4  | 559 | 1  | 41  | 4  | 370 | 2  | 216 |
| Slic17a7 - layer 4  | 4  | 873 | 4  | 1061 | 1  | 154 |
| Slic17a7 - layer 5  | 4  | 806 | 3  | 201 | 3  | 342 |
| Cux2 - layer 2/3 | 7  | 1928 | 5  | 875 | 6  | 1127 | 7  | 770 | 4  | 504 | 5  | 379 |
| Cux2 - layer 4  | 7  | 1954 | 3  | 915 | 3  | 704 | 3  | 614 | 4  | 789 | 3  | 214 |
| Rorb           | 4  | 984 | 2  | 513 | 5  | 896 | 6  | 805 | 5  | 475 | 4  | 807 |
| Scnn1a         | 5  | 881 |
| Nfl5a1         | 6  | 350 | 5  | 421 | 4  | 177 | 5  | 182 | 3  | 80  | 4  | 942 |
| Rbp4           | 6  | 325 | 4  | 272 | 4  | 306 | 6  | 397 | 6  | 254 | 2  | 47 |
| Fezf2          | 3  | 290 | 2  | 257 |
| Tlx3           | 1  | 303 |
| Ntsr1          | 3  | 230 | 4  | 457 | 4  | 355 |
| Sat - layer 4  | 4  | 49  | 6  | 128 | 1  | 24  | 5  | 94  |
| Sat - layer 5  | 6  | 56  | 5  | 65  | 3  | 58  |
| Vip - layer 2/3 | 4  | 77  | 5  | 88  | 5  | 133 |
| Vip - layer 4  | 6  | 115 | 4  | 46  | 7  | 136 |

**Extended Data Fig. 8 | Populations for running correlation analysis.** Table summarizing the number of experiments and neurons, for each Cre line, layer, area combination, included in the running correlation analysis. These are from sessions in which the mouse was running between 20–80% of the time.
**Wavelet models**

| Cre      | V1 expts | V1 neurons | LM expts | LM neurons | AL expts | AL neurons | PM expts | PM neurons | AM expts | AM neurons | RL expts | RL neurons |
|----------|----------|------------|----------|------------|----------|------------|----------|------------|----------|------------|----------|------------|
| Emx1 - layer 2/3 | 3        | 296        | 4        | 217        | 2        | 160        | 2        | 127        | 1        | 28         | 3        | 178        |
| Emx1 - layer 4   | 2        | 247        | 2        | 152        |          |            | 1        | 86         | 1        | 36         | 2        | 168        |
| Emx1 - layer 5   | 4        | 324        |          |            | 3        | 136        | 1        | 31         | 1        | 5          | 1        | 59         |
| Slic17a7 - layer 2/3 | 5        | 366        | 6        | 274        | 1        | 14         | 5        | 136        | 2        | 66         | 1        | 14         |
| Slic17a7 - layer 4 | 4        | 446        | 2        | 207        | 1        | 123        | 4        | 124        |          |            |          |            |
| Slic17a7 - layer 5 | 3        | 200        | 5        | 168        |          |            | 6        | 290        | 1        | 23         |          |            |
| Cux2 - layer 2/3 | 8        | 769        | 6        | 431        | 6        | 362        | 9        | 350        | 5        | 160        | 6        | 204        |
| Cux2 - layer 4   | 6        | 639        | 4        | 359        | 7        | 663        | 4        | 270        | 6        | 236        | 4        | 180        |
| Rorb              | 7        | 652        | 6        | 398        | 6        | 376        | 7        | 242        | 8        | 191        | 2        | 114        |
| Scnn1a            | 9        | 531        |          |            |          |            |          |            |          |            |          |            |
| Nr5a1             | 8        | 236        | 6        | 122        | 6        | 59         | 7        | 123        | 6        | 68         | 5        | 401        |
| Rbp4              | 7        | 151        | 7        | 177        | 6        | 123        | 6        | 209        | 8        | 129        | 4        | 38         |
| Fezf2             | 4        | 177        | 5        | 237        |          |            |          |            |          |            |          |            |
| Tlx3              | 5        | 400        | 2        | 209        |          |            |          |            |          |            |          |            |
| Ntr1              | 6        | 206        | 7        | 186        |          |            |          |            |          |            |          |            |
| Sst - layer 4     | 9        | 64         | 8        | 65         | 1        | 10         | 6        | 49         | 1        | 5          |          |            |
| Sst - layer 5     | 7        | 34         | 6        | 30         |          |            | 8        | 35         | 1        | 4          |          |            |
| Vip - layer 2/3   | 9        | 80         | 8        | 91         |          |            | 7        | 68         |          |            |          |            |
| Vip - layer 4     | 7        | 34         | 9        | 59         |          |            |          |            | 9        | 74         |          |            |

**Extended Data Fig. 9 | Populations for wavelet model analysis.** Table summarizing the number of experiments and neurons for each Cre line, layer, area combination for which wavelet models were fit. The neurons had to be present in all three imaging sessions to be included.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  Give P values as exact values whenever suitable.

- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson’s r), indicating how they were calculated
  Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection
Two photon imaging data was collected using either Nikon Elements (v4.3) or Sciscan (v12.0), with custom Python 2.7 scripts to run the workflow and interface with the eye-tracking and behavior cameras. Intrinsic signal imaging data was collected using custom scripts written in Python.

Data analysis
All analyses were performed using custom scripts written in Python 2.7, using NumPy, SciPy, Pandas, Matplotlib, Seaborn, Keras, and Tensorflow, or MATLAB. Analysis code is available at http://alleninstitute.github.io/AllenSDK/ and https://github.com/alleninstitute/visual_coding_2p_analysis. Event extraction was performed using FastLZeroSpikeInference available at https://github.com/jewellsean/FastLZeroSpikeInference.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
  - Accession codes, unique identifiers, or web links for publicly available datasets
  - A list of figures that have associated raw data
  - A description of any restrictions on data availability

This is an openly available dataset, accessible via a dedicated web portal (http://observatory.brain-map.org/visualcoding), with a custom Python-based Application Programming Interface (API), the AllenSDK (http://alleninstitute.github.io/AllenSDK/).
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences  ☐ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample size was determined qualitatively to balance repeated experiments for each area/layer/Cre line combination and the preserve the breadth of survey. Our sample size matched, or exceeded, those found in previous publications.

Data exclusions

Mice were excluded for evidence of epileptiform activity, and individual imaging sessions were failed if there were signs of bleaching, saturation, excessive z-drift, or animal stress, among other factors, as described in our Methods.

Replication

Data acquired from multiple mice from multiple litters per transgenic line surveyed. Extensive acquisition metadata as well as detailed white papers are reported as part of the Allen Cell Brain Observatory (observatory.brain-map.org); these additional details are intended to aid other laboratories if they seek to replicate the results presented in this study.

Randomization

Within each transgenic Cre line, mice were randomly assigned to data collection in order to sample different areas and imaging depths. Stimulus conditions for gratings and natural scenes were presented in a randomized order within each epoch, as described. No other randomization was used as there were fixed experimental condition for all other aspects of the data set.

Blinding

Blinding was not relevant to this study as there was a single experimental condition for all data collected.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| Involved in the study |
|-----------------------|
| ☑ Antibodies          |
| ☑ Eukaryotic cell lines |
| ☑ Palaeontology       |
| ☑ Animals and other organisms |
| ☑ Human research participants |
| ☑ Clinical data       |

Methods

| Involved in the study |
|-----------------------|
| ☑ ChIP-seq            |
| ☑ Flow cytometry      |
| ☑ MRI-based neuroimaging |

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

| Laboratory animals |
|-------------------|
| Mus musculus, male and female, mean age 108 ± 17 (st. dev.) days Wild animals |

| Wild animals |
|--------------|
| This study did not use wild animals. |

| Field-collected samples |
|-------------------------|
| This study did not use samples collected from the field |

| Ethics oversight |
|------------------|
| Experiments involving mice were approved by the Institutional Animal Care and Use Committees of the Allen Institute for Brain Science in accordance with NIH guidelines. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.